

Tecnologías de la Traducción: Proyecto colaborativo de traducción*

Felipe Sánchez Martínez, Mikel L. Forcada

Departament de Llenguatges i Sistemes Informàtics
Universitat d'Alacant

Curso 2016–2017, segundo cuatrimestre
última actualización: 11 de octubre de 2017

1. Introducción

Este documento recoge la especificación del proyecto colaborativo de traducción que tenéis que hacer en casa y durante las sesiones de laboratorio L_{13} – L_{21} en grupos de 4 personas.

El objetivo del proyecto colaborativo de traducción es que pongáis en práctica todo aquello que habéis aprendido durante la realización de las actividades prácticas (véase el documento “Tecnologías de la Traducción: Actividades prácticas”) en un proyecto de traducción real, aprendáis a evaluar la utilidad de las tecnologías de la traducción disponibles (traducción automática y memorias de traducción) para un uso concreto, a elegir cuáles usaréis y a usarlas conjuntamente para la realización del proyecto. Para esto debéis proponer un proyecto de traducción al profesorado para que os dé el visto bueno (opción preferible), o escoger un proyecto de entre los que el profesorado pondrá a vuestra disposición en UACloud (véase la sección 3). La lengua meta de la traducción tiene que ser español o catalán.

Importante: Tenéis que proponer el proyecto que abordaréis, encontrar los ficheros a traducir (si hace falta), conformar el grupo de 4 personas que trabajaréis conjuntamente y rellenar el documento `tt-projectes-grups-es.odt` que encontraréis

*© 2016–2017 Universitat d'Alacant. Este material puede ser distribuido, copiado y exhibido si los nombres de los autores se muestran en los créditos. Las obras derivadas tienen que distribuirse bajo los mismos términos de licencia que el trabajo original. Más detalles: <http://creativecommons.org/licenses/by-sa/3.0/deed.es>. Podéis pedir los fuentes LaTeX a los autores.

Entrega de propuestas	2 de abril de 2017
Aprobación de propuestas	5 de abril de 2017
Propuesta definitiva	9 de abril de 2017
Ejecución del proyecto	Sesiones L_{13} – L_{21}
Entrega 1a convocatoria	6 de junio de 2017
Entrega 2a convocatoria	11 de julio de 2017

Tabla 1: Calendario-resumen de las fechas relacionadas con la propuesta de proyectos, sesiones de laboratorio que se dedicarán y fecha de entrega del proyecto en cada convocatoria.

en la sección de materiales de UACloud. Este documento lo tiene que entregar uno de los miembros del grupo mediante la aplicación de UACloud *Evaluación* antes del 2 de abril de 2017. El profesorado estudiará la propuesta y os dará el visto bueno antes del 5 de abril de 2017. Si hace falta que hagáis alguna modificación a vuestra propuesta, la tendréis que hacer antes del 9 de abril de 2017 (véase el calendario-resumen en la tabla 1). Tened cuidado y no os dejéis esto para el último momento.

Advertencia: Quien no haya entregado o no figure en ninguna propuesta de proyecto de traducción antes de la fecha tope establecida no podrá trabajar en ningún proyecto de traducción y tendrá una calificación de *cero* en esta parte de la asignatura.

Advertencia: Si la propuesta de proyecto no tiene la calidad mínima exigible, el profesorado os podrá asignar

un proyecto a su elección.

2. Evaluación

El proyecto colaborativo de traducción servirá para establecer el 24 % de la nota global de la asignatura (véase el documento “Tecnologías de la Traducción: Información y normas del curso”). La fecha límite para entregarlo es el 6 de junio de 2017 a las 23.59 horas. La entrega se tiene que hacer mediante el portafolios electrónico, adonde tenéis que subir el informe del proyecto y enlazarlo desde el documento maestro.

Como que el proyecto colaborativo de traducción lo haréis en grupo, tenéis que decidir en qué portafolios entregaréis el proyecto e indicarlo en vuestra propuesta.

Si la asignatura se suspende en junio y hay que repetir el proyecto colaborativo de traducción, se volverá a calificar lo que hayáis entregado el 11 de julio de 2017 a las 23.59.

3. Proyectos

En UACloud, en la carpeta *Proyectos / Proyectos* de la sección de Materiales, encontraréis los ficheros correspondientes a cada uno de los proyectos colaborativos de traducción propuestos por el profesorado. Dentro de cada fichero comprimido encontraréis una descripción del proyecto, recursos como por ejemplo memorias de traducción o traducciones automáticas y, si es necesario, instrucciones sobre como llevar a cabo el proyecto o localizar los ficheros a traducir.

Como ya se ha explicado más arriba, es preferible que propongáis vosotros un proyecto de traducción al profesorado; en este caso, conviene que habléis con vuestro profesor de prácticas antes de la preparación de la propuesta.

El profesorado valorará muy positivamente que propongáis vuestro propio proyecto de traducción y que éste sea un proyecto de traducción real.

Lengua	Número de palabras
Alemán	1.700
Inglés	2.100
Francés	2.600
Italiano	3.000
Catalán	3.500
Español	3.500

Tabla 2: Número de palabras a traducir **por miembro del grupo** en función de cual sea la lengua origen de la traducción. Recordad que la lengua meta tiene que ser español o catalán.

Advertencia: A la hora de proponer un proyecto de traducción aseguraos que los documentos no se encuentran ya traducidos en Internet.

Advertencia: Aseguraos de que los textos a traducir están en un formato compatible con OmegaT: .odt, .doc, .docx, .html, .rtf, etc. **PDF no** es un formato compatible con OmegaT.

Si vais a traducir documentos HTML de un sitio web, podéis usar la herramienta *HTTrack* para descargarlos los documentos a traducir. Esta herramienta es libre/de código fuente abierto y la podéis descargar desde la URI <https://www.httrack.com/>.

Cantidad de texto. El número de palabras que tiene que traducir cada miembro del grupo está en función del par de lenguas concreto. En la tabla 2 encontraréis la cantidad mínima de palabras a traducir por miembro del grupo en función de cuál sea la lengua origen. Cuando preparéis vuestra propuesta deberéis tener en cuenta esta información y el hecho de que se espera que dediquéis al proyecto 18 horas de trabajo en el laboratorio y alrededor de 27 horas en casa, todo esto por miembro, claro. Estudiada vuestra propuesta el profesorado podrá pedir os que traduzcáis una cantidad de texto diferente.

4. Ejecución del proyecto

Para llevar a cabo vuestro proyecto de traducción podéis usar cualquier recurso que os sea de ayuda. En el apéndice A encontraréis una lista no exhaustiva de recursos que pueden ser de vuestro interés.

La ejecución del proyecto estará dividida en las siguientes tareas:

1. Planificación del proyecto (sección 4.1).
2. Estimación de los costes de traducción sin el apoyo de ninguna tecnología (sección 4.2).
3. Estimación de los costes de traducción usando traducción automática y postedición (sección 4.3).
4. Estimación de los costes de traducción usando memorias de traducción (sección 4.4).
5. Decisión sobre la adopción de tecnologías y estimación del coste total de ejecución del proyecto (sección 4.5)
6. Traducción mediante OmegaT (sección 4.6).
7. Control de calidad (sección 4.7).

Para poder abordar una tarea con garantías de éxito tenéis que haber finalizado la tarea anterior. Además de ejecutar las tareas mencionadas más arriba, tenéis que escribir un informe (véase la sección 5).

4.1. Planificación del proyecto

La primera tarea del proyecto consiste a hacer una planificación semanal del trabajo a hacer a lo largo de las 4 semanas que durará la ejecución del proyecto. Para lo que tenéis que leer antes el resto de este documento para tener una visión general del trabajo que tenéis que hacer.

Para hacer la planificación del proyecto tenéis que tener en cuenta lo siguiente:

- La sesión de laboratorio L_{13} se dedicará a resolver dudas sobre la planificación del proyecto y sobre la búsqueda de recursos que os sean de utilidad (véase el apéndice A) y a explicar algunos conceptos sobre traducción automática estadística.
- La sesión de laboratorio L_{14} se dedicará a discutir como se tienen que hacer las estimaciones de coste y a resolver dudas sobre la estimación del coste de traducción sin el apoyo de ninguna tecnología.
- La sesión de laboratorio L_{15} se dedicará a resolver dudas sobre la estimación del coste de traducción mediante postedición de la traducción automática.

- La sesión de laboratorio L_{16} se dedicará a resolver dudas sobre la estimación del coste de traducción con el apoyo de memorias de traducción.
- La primera parte de la sesión L_{17} se dedicará a resolver dudas sobre la estimación del coste global de traducción del proyecto.
- La segunda parte de la sesión de laboratorio L_{17} y la sesión de laboratorio L_{18} se dedicarán a la traducción de los documentos del proyecto y a la resolución de dudas.
- La sesión de laboratorio L_{19} se dedicará al control de calidad.
- El resto de sesiones de laboratorio se dedicarán a la resolución de dudas.

Consultad la tabla 3 del documento “Tecnologías de la Traducción: información y normas del curso” para conocer, para cada grupo de prácticas, la fecha concreta de las sesiones mencionadas más arriba.

En vuestra planificación tenéis que tener en cuenta que además de traducir tenéis que escribir un informe que tenéis que entregar al final.

Es importante que introduzcáis la escritura del informe en vuestra planificación semanal y que vayáis escribiéndolo según que vais ejecutando cada una de las tareas.

En el informe tendréis que describir como se ha desarrollado el proyecto; para lo cual, tendréis que hacer una planificación exhaustiva por adelantado y hacer un seguimiento de la misma.

4.2. Estimación del coste de traducción sin la ayuda de ninguna tecnología

Durante todo el proyecto nos referiremos a un cliente hipotético que es quien ha encargado la traducción de vuestro proyecto, a los costes de traducción, y a los precios que negociaríais con este cliente.

Ante todo, vamos a ver como podemos estimar los costes de traducción si no tenemos el apoyo de ninguna tecnología. Esta estimación nos servirá de referencia a la hora de decidir qué tecnologías adoptamos y cuáles no.

¿Cuánto querríais ganar traduciendo? Ya os podéis imaginar que una buena parte, digamos un 30 %, de lo que cobráis a vuestro cliente se irá en impuestos y gastos y a vosotros os quedará un 70 % aproximadamente (esta estimación está hecha muy por encima y habría que hacer un estudio económico detallado que cae fuera del alcance del proyecto). Si queréis ganar, por ejemplo, unos 3 000 € brutos al mes, esto quiere decir que tendréis que cobrar

$$3\,000 \text{ €/mes} \times \frac{100\%}{70\%} \simeq 4\,200 \text{ €/mes.}$$

Si trabajáis unas 140 horas al mes (es decir, unas 8 horas por día laborable), cobraréis unos 30 € por cada hora de vuestro trabajo.

¿Cuál sería un precio razonable por palabra? Se suele decir que el ritmo profesional de traducción es de unas 2 500 palabras por día, es decir, unos $2\,500/8 \simeq 300$ palabras por hora. Las 300 palabras por hora cuestan a vuestro cliente 30 €. Por lo tanto, el precio por palabra que paga viene a ser aproximadamente

$$\frac{30 \text{ €/hora}}{300 \text{ palabras/hora}} = 0,10 \text{ €/palabra.}$$

¿Todo el mundo traduce igual de rápido? Claramente no. Depende de quién traduce, de cuál es el par de lenguas, de la naturaleza del texto concreto, etc. ¿Cómo podéis saberlo? Tenéis que traducir una muestra de texto y medir el tiempo, es decir, *medir vuestra productividad*: calcular cuántas palabras traducís en una cierta cantidad de tiempo, o cuánto tiempo tardáis en traducir una cierta cantidad de palabras.

Ejemplo: Habéis tardado 1 hora y 22 minutos, es decir, $1 + 22/60 = 1,37$ horas, en traducir 500 palabras. Vuestra productividad es por lo tanto

$$\frac{500 \text{ palabras}}{1,37 \text{ horas}} \simeq 365 \text{ palabras/hora.}$$

La muestra tiene que ser *representativa* del encargo —en vuestro caso, partes cuidadosamente seleccionadas del encargo mismo— y suficientemente *grande* (véase más abajo) para extraer conclusiones que sean extrapolables al resto de texto de vuestro proyecto.

Traducid directamente desde OmegaT (creando un pequeño proyecto de traducción sin memorias ni apoyo

de la traducción automática): cargad la muestra de texto a traducir y cronometráos, evitando interrupciones, claro está. Como resultado, obtendréis una memoria de traducción con segmentos ya traducidos que podréis usar más adelante para no tener que volver a traducirlos.

Cuando, haciendo una regla de tres, sepáis cuántas palabras traducís por hora, podéis calcular vuestro precio por palabra, que denominaremos *precio completo (full rate)*:

Ejemplo: Traducís unas 365 palabras/hora y queréis cobrar la hora a 30 €. El precio por palabra es por lo tanto

$$\frac{30 \text{ €/hora}}{365 \text{ palabras/hora}} \simeq 0,08 \text{ €/palabra.}$$

Advertencia: Debéis estimar cuantas palabras traduce por hora cada miembro del grupo por separado y todos juntos. Para calcular cuantas palabras traducís por hora como grupo **no** tenéis que hacer la media de la productividad calculada para cada miembro; tenéis que hacer el cálculo sumando, por un lado, la cantidad de palabras traducidas entre todos y, por otro lado, el tiempo invertido entre todos en la traducción.

¿Cuánto texto? Una manera de saber si la cantidad de texto que habéis usado para hacer la estimación es suficiente es estudiar, por ejemplo, si vuestra productividad varía mucho de la primera mitad del texto a la segunda mitad. Si varía mucho quiere decir que la cantidad de texto que habéis usado no es suficiente para tomar una decisión; quizás tenéis que usar más texto para estimar mejor el coste.

¿Lo habéis medido bien? El hecho de que los resultados varíen también tiene implicaciones a la hora de presentar los resultados. Por ejemplo, no tiene sentido decir que traducís 362 palabras/hora (con tanta exactitud) cuando para la primera mitad del texto habéis traducido 358 palabras/hora y para la segunda 363 palabras/hora. Quizás sea más lógico decir “alrededor de 360 palabras/hora”.

4.3. Estimación del coste de traducción con la ayuda de la traducción automática

En esta tarea tenéis que elegir un sistema de traducción automática accesible por Internet para el par de lenguas de vuestro proyecto y hacer *una estimación del coste de traducción* mediante la postedición de las traducciones en bruto que produce para este par de lenguas. Esto evaluará la *calidad* del sistema para el proyecto mediante una estimación de los costes de traducción —en comparación con los costes de traducción sin el apoyo de ninguna tecnología— y servirá para decidir sobre la *adopción* del traductor automático. Para lo cual, *posteditaréis* una parte de los textos de vuestro proyecto; podéis ver como usar traducción automática desde dentro de OmegaT en el apéndice B. Si hubiera más de un sistema de traducción automática para el par de lenguas en cuestión, también podríais hacer una evaluación comparativa entre los sistemas.

Como en el apartado 4.2, para hacer la evaluación tenéis que elegir los textos de forma que sean suficientemente grandes y representativos de toda la tarea de traducción de vuestro proyecto. Hacedlo con otra parte del texto (así iréis avanzando en el desarrollo del proyecto).

Sobre la postedición La *postedición* es la modificación *mínima* de una traducción generada por ordenador para *hacerla adecuada a un propósito bien definido*; por lo tanto, no hay que hacer cambios *preferenciales* (“esta solución adecuada *me gusta más* que esta otra que también es adecuada”). Los cambios estilísticos se tienen que hacer estrictamente cuando, si no se hicieran, las traducciones de vuestro proyecto no cumplirían con el propósito para el cual fueron encargadas. Las modificaciones pueden ser: *borrados* de una palabra que sobra, *sustituciones* de una palabra por otra, o *inserciones* de una palabra que falta. Tienen que ser las *mínimas* necesarias: si hay más de una edición posible, elegid la que se haga con el mínimo de modificaciones necesarias. Por un lado, esto permite estimar mejor la utilidad real de la traducción en bruto; por otro lado, a la larga, os ahorraréis trabajo si aprendéis a trabajar así.

¿Me ayuda la traducción automática? Para saber si os ayuda la traducción automática, tenéis que repetir el estudio de productividad del apartado 4.2, pero posteditando el resultado de la traducción automática en vez de

traducir desde cero. Pueden pasar tres cosas:

1. Que la traducción automática no os ayude y, de hecho, os haga perder tiempos (por ejemplo, en vez de las 365 palabras/hora que en el ejemplo de la sección 4.2 hacíais a mano, traducís 295 palabras/hora). Olvidaos de ella.
2. Que la traducción automática ni os ayude ni os perjudique (por ejemplo, traducís unas 360 palabras/hora). Quizás sea buena idea olvidaros también.
3. Que la traducción automática os ayude y vuestra productividad aumente (por ejemplo, llegáis a traducir 480 palabras/hora). En este caso, el coste de vuestra traducción será más bajo.

Ejemplo: Traducís unas 480 palabras/hora y queréis cobrar, como antes, la hora a 30 €. El precio por palabra es ahora más barato que el *precio completo* y permite ajustar mejor el presupuesto que presentáis a vuestro cliente:

$$\frac{30 \text{ €/hora}}{480 \text{ palabras/hora}} \simeq 0,06 \text{ €/palabra.}$$

No os fiéis de vuestras impresiones sobre la utilidad de la ayuda que ofrece la traducción automática¹ en cuanto a la productividad: medíos bien para saber si ganáis o perdéis dinero.

Como cuando habéis estimado el coste sin el apoyo de ninguna tecnología de traducción (apartado 4.2), posteditad los textos directamente desde OmegaT para obtener una memoria de traducción con segmentos ya traducidos que podréis usar más adelante.

Indicadores indirectos de la utilidad de la traducción automática: A veces es difícil medir bien el tiempo (por interrupciones, etc.). Es buena idea tener también una idea de la tasa de postedición por palabra, es decir, cuántas palabras habéis borrado, sustituido o insertado por cada 100 palabras, expresado como un porcentaje. Hay programas que calculan esto automáticamente después de hacer el trabajo.² Más concretamente, la tasa de postedición es la cantidad *mínima* de palabras que hay que borrar, sustituir o insertar para convertir la traducción en bruto en una traducción adecuada *para el*

¹Son engañosas: mirad el estudio <http://langtech.autodesk.com/productivity.html>

²Por ejemplo, <http://xixona.dlsi.ua.es/~fran/eval/>

propósito previsto, expresada como porcentaje del número total de palabras de la traducción posteditada.³

Son posibles otros indicadores cuantitativos de la calidad, algunos de los cuales son mejoras de la tasa de postedición mencionada más arriba. Tened en cuenta que no cuesta lo mismo corregir todos los errores cometidos por el traductor automático —hay errores que son más fáciles de corregir que otros— y la tasa de postedición por palabra no tiene en cuenta la dificultad de cada error que habéis corregido.

Comparando la estimación de costes para varios proyectos, es posible estudiar como varía vuestro tiempo de postedición con la tasa de postedición por palabra, lo que os permitirá hacer el próximo presupuesto sin tener que ir con un cronómetro: usaréis la tasa de postedición como indicador cualitativo *indirecto* de la utilidad (es decir, de la calidad) de la traducción automática.

Apartado optativo: Podéis hacer un estudio que clasifique los errores en categorías en cuanto al impacto en la productividad de postedición e incluir en vuestro informe una breve discusión.

4.4. Estimación de los costes de traducción con la ayuda de memorias de traducción

Las memorias de traducción pueden mejorar mucho vuestra productividad: si hay segmentos de vuestro proyecto que ya han sido traducidos totalmente o parcialmente, vuestro trabajo consiste a comprobar qué parte está ya traducida y si está bien traducida, y traducir las partes que no están traducidas, pero antes de decidir si las usaremos o no, tenemos que evaluar si de verdad nos ahorran trabajo.

¿Cuánto ayudan las memorias de traducción? Para saber cuánto os ayudan las memorias de traducción —si es que os ayudan—, tenéis que repetir de nuevo el estudio de productividad del apartado 4.2, pero partiendo de las propuestas de vuestra memoria de traducción cuando las haya. Podéis calcular vuestro precio por palabra igualmente, pero es posible que vuestro cliente sea quién os proporcione la memoria y quiera negociar el precio

por franjas de porcentaje de concordancia (*fuzzy match*) cómo es costumbre en el mercado.

Ejemplo: Imaginad que acabáis poniéndoos de acuerdo sobre el siguiente esquema de precios:⁴

Porcentaje de concordancia	Porcentaje del precio completo por palabra
95 %–100 %	10 %
85 %–94 %	30 %
75 %–84 %	50 %
por debajo del 75 %	100 %

Si vuestro *precio completo* es 0,08 €/palabra, los precios de cada franja, aplicando los porcentajes, serían

Porcentaje de concordancia	Precio por palabra
95 %–100 %	0,008 €
85 %–94 %	0,024 €
75 %–84 %	0,040 €
por debajo del 75 %	0,080 €

Si habéis negociado bien, el precio total del trabajo calculado multiplicando el número de palabras en cada franja por el precio correspondiente y el precio total calculado usando la productividad media en palabras por hora no se tendrían que diferenciar demasiado.

Para calcular el coste prácticamente todas las herramientas existentes de traducción asistida os indican cuántas palabras de vuestro proyecto tenéis en cada una de las franjas.

Análisis de coincidencias con OmegaT: Para obtener información detallada sobre los segmentos y las coincidencias encontradas en las memorias de traducción que vais a usar tenéis que ir al menú Herramientas→Estadísticas de coincidencias.

Advertencia: Las estadísticas sobre coincidencias las tenéis que obtener **antes de empezar a traducir**. En caso contrario, los datos obtenidos no serán fieles a la realidad.

³A veces, este indicador se denomina *word error rate* o tasa de error por palabra.

⁴Podéis encontrar otros esquemas en webs como por ejemplo <http://proz.com>.

Advertencia: Si estáis trabajando con memorias muy grandes, antes de iniciar el programa OmegaT tenéis que abrir con la *Libreta o Bloc de notas* el fichero OmegaT.bat que estará alojado en la carpeta donde descomprimisteis OmegaT y cambiar -Xmx512M por -Xmx1024M para permitir que OmegaT use hasta 1024 MB de memoria RAM. Para que este cambio tenga efecto tenéis que arrancar OmegaT haciendo doble clic sobre el fichero OmegaT.bat (en lugar de OmegaT.jar). Si vuestro ordenador tiene un sistema operativo GNU/Linux o MacOS, tenéis que hacer lo mismo pero con el fichero OmegaT.

Cuando tengáis las estadísticas de coincidencias, las podéis usar para estimar los costes.

Ejemplo: Supongamos que las estadísticas sobre coincidencias que nos da OmegaT son:

	Segmentos	Palabras
95 %-100 %:	11	105
85 %-94 %:	15	184
75 %-84 %:	13	183
50 %-74 %:	17	260
Sin coine.:	13	120

De acuerdo con los precios por franjas que se mencionan más arriba, el precio total de vuestro proyecto sería:

$$\begin{array}{rcl}
 105 \times 0,008 & + & \\
 + 184 \times 0,024 & + & \\
 + 183 \times 0,04 & + & \\
 + 260 \times 0,08 & + & \\
 + 120 \times 0,08 & = & 42,97\text{€}.
 \end{array}$$

En cambio, traduciéndolo todo a mano el coste total sería

$$(105+184+183+260+120) \times 0,08 = 68,16\text{€}.$$

En este caso, el precio de la traducción, de acuerdo con lo que se ha negociado, bajaría aproximadamente un 37 %.

En el caso de que no obtengáis muchas coincidencias de vuestras memorias, el coste de traducción no bajaría apreciablemente del coste sin apoyo tecnológico.

Fijaos que comparando la estimación de costes para varios proyectos, es posible estudiar como varía vuestra productividad con las estadísticas de coincidencias en cada franja, cosa que os permitirá negociar mejor los precios en cada franja con clientes futuros.

4.5. Decisión sobre la adopción de tecnologías y estimación del coste total de ejecución del proyecto

Si la traducción automática ayudaba y las memorias también, os puede convenir combinarlas. Si el precio que habéis calculado para traducción automática posteditada en el ejemplo era 0,06 €/palabra, está claro que podéis mejorar de 0,08 €/palabra a 0,06 €/palabra el coste de la franja “por debajo del 75 %” usando traducción automática para esta franja y las memorias para el resto de las franjas.

Con sistemas de traducción automática mejores, podría ser que el precio por palabra de postedición fuera de 0,03 €/palabra y entonces podríais descartar también las sugerencias de la memoria en la franja 75 %-84 %.

Las estimaciones sirven para negociar: ahora ya sabéis estimar cuánto tendríais que pagar vuestro cliente para que podáis ganaros el jornal que esperabais. De ahora en adelante, depende de vuestra habilidad para negociar un buen precio con el cliente.

Para saber más...

- Un web interesante para calcular la compensación justa por vuestro trabajo es el web <http://fairtradetranslation.org>.
- La asociación TAUS publicó unas directrices para fijar el precio de las traducciones cuando se postedita la traducción automática: <http://is.gd/tauspricing>.

4.6. Traducción mediante OmegaT

Cuando hayáis estimado los costes de traducción tenéis que traducir los documentos del proyecto mediante el software de asistencia a la traducción OmegaT.

Los textos que hayáis traducido para recoger los datos que os han permitido estimar los costes de traducción del proyecto los podéis usar como memoria de traducción para evitar volver a traducirlos. Si no habéis usado OmegaT para la postedición hecha para evaluar el traductor automático o a la hora de hacer un pequeño trabajo de traducción, podéis generar una memoria de traducción con alguno de los servicios de alineamiento mencionado en el apéndice A.

Advertencia: Aunque no uséis memorias de traducción, el uso de OmegaT es **obligatorio**. En el apéndice B se explica como usar traducción automática desde OmegaT.


Advertencia: Recordad que tenéis que tener activados el corrector ortográfico y el corrector gramatical en OmegaT, tal como se explica en la sección correspondiente del documento “Tecnologías de la Traducción: Actividades prácticas”.

Advertencia: Esto lo tiene que hacer un único miembro del grupo, una única vez.

Cuando hayáis acabado de incorporar todos los ficheros a la copia local del proyecto de traducción, cada miembro del grupo se hará una copia de ésta en su ordenador. De aquí en adelante cada uno trabajará con su copia local. OmegaT enviará las traducciones que vayáis haciendo al servidor automáticamente cada 3 minutos, y se descargará las que hayan hecho vuestros compañeros, de forma que podréis usarlas si las encontráis de utilidad.

4.6.1. Trabajo en grupo con OmegaT

Para que podáis trabajar en grupo y compartir las unidades de traducción que iréis creando durante la traducción, cada grupo tiene a su disposición un espacio para el proyecto de traducción en un servidor del Departament de Llenguatges i Sistemes Informàtics. Para trabajar con este proyecto uno de los miembros del grupo tiene que hacer lo siguiente:

1. Abrir OmegaT y seleccionar la opción *Descargar proyecto en equipo* del menú *Proyecto*; se abrirá un diálogo.
2. En *URL del dispositivo* introducir la dirección de vuestro proyecto: `https://sicvs.dlsi.ua.es/tt/X` donde *X* se el DNI/NIE que elegisteis para entregar el proyecto colaborativo.
3. Hacer clic en el botón  y seleccionar la carpeta local (del disco duro o de la memoria USB) donde se guardará el proyecto de traducción.
4. Hacer clic en el botón *Aceptar*; se abrirá un diálogo para que introduzcáis vuestras credenciales.
5. Introducir vuestro usuario (DNI/NIE que elegisteis para entregar el proyecto), vuestra contraseña (la misma que usáis para acceder al portafolios) y hacer clic en el botón *Aceptar*.

Si todo ha ido bien tendréis una copia local del proyecto de traducción en el ordenador de quien ha hecho los pasos mencionados arriba. Ahora tenéis que importar los ficheros a traducir y poner las memorias de traducción y los glosarios que usaréis en las carpetas *tm* y *glossary*, respectivamente, de esta copia local del proyecto de traducción. Estos ficheros no se copiarán al servidor.

4.6.2. Sobre el formato de los documentos a traducir

OmegaT esconde la información de formato de los ficheros a traducir usando etiquetas sencillas del estilo de las de XML; tened cuidado de no quitar o añadir ninguna.

Para comprobar si por error habéis borrado alguna de las etiquetas que OmegaT utiliza para preservar el formato del texto traducido tenéis que seleccionar la opción del menú *Herramientas*→*Validar etiquetas*. OmegaT os mostrará una ventana con los segmentos donde ha encontrado que faltan etiquetas.

4.7. Control de calidad

Cuando hayáis acabado de traducir todos los ficheros de vuestro proyecto tenéis que llevar a cabo un control de la calidad de la traducción resultante. El control de calidad consiste en hacer:

- Una revisión de la traducción para detectar posibles errores introducidos durante el proceso de traducción. La revisión la tiene que hacer un miembro del grupo que no haya traducido el fragmento de texto que esté revisando.
- Una revisión monolingüe del texto ya traducido, sin tener acceso al texto fuente. La finalidad de esta revisión es garantizar la consistencia terminológica y de estilo en todo el documento.
- Opcionalmente, podéis llegar a acuerdos con otros grupos para hacer una evaluación cruzada de la calidad de la traducción (es decir, los miembros de un grupo hacen el control de calidad de las traducciones hechas por otro grupo y viceversa). Si lo hacéis, indicadlo en el informe.

Cómo veis, os tenéis que dividir el trabajo, a pesar de que todos los miembros del grupo tienen que traducir, revisar la traducción hecha por otro miembro del grupo y hacer parte de la revisión monolingüe del texto meta.

Advertencia: Las correcciones que podáis hacer como resultado del control de calidad las tenéis que hacer mediante OmegaT, no directamente en el documento traducido, para que las memorias de traducción resultantes queden también corregidas.

La información que se tiene que añadir al informe como resultado de cada una de las tareas del proyecto se especifica en la plantilla de informe.

Se valorará mucho la *claridad* y la *brevedad* del informe; no se aceptarán informes sin un nivel mínimo de corrección lingüística y de estilo o que no sigan la plantilla que el profesorado ha puesto a vuestra disposición.

4.7.1. Comprobaciones automáticas de la calidad

OmegaT puede hacer algunas comprobaciones y ayudaros a detectar algunos errores de traducción de forma automática. Para esto tenéis que seleccionar la opción Programación... del menú Herramientas; se abrirá una ventana como la que se muestra en la figura 1. En esta ventana tenéis que seleccionar la opción QA – Check Rules del panel de la izquierda y hacer clic en el botón Ejecutar que encontraréis en la parte inferior de la ventana. Si en el panel de la izquierda no aparecen varias opciones como las que se muestran a la figura, tenéis que seleccionar la carpeta donde están los *programas*: la carpeta con nombre *scripts* que encontraréis en vuestra instalación de OmegaT.

Una vez se haya ejecutado el programa se abrirá otra ventana con los errores detectados; podéis hacer clic sobre el número de segmento para ir directamente al segmento donde se encuentra el error y corregirlo. También podéis elegir las comprobaciones a hacer y volver a ejecutar el programa haciendo clic sobre el botón Refresh; fijaos que por defecto no comprueba si hay segmentos con errores ortográficos.

5. Estructura del informe

El informe que publicaréis en vuestro portafolios *web* tiene que tener un estilo objetivo y profesional y tiene que estar escrito de forma que una persona ajena a la asignatura lo pueda entender y pueda reproducir vuestro trabajo, independientemente de este enunciado. En la sección de materiales encontraréis la plantilla de informe que tenéis que utilizar. Descargad el archivo *ttpro-es.zip* y descomprimidlo para obtener el fichero *ttpro.html* con la plantilla de informe que tenéis que utilizar.

Requisitos técnicos que tiene que cumplir el proyecto para ser evaluado (si no se cumple alguno de los requisitos, la calificación será *cero*):

- Hay un informe llamado *ttpro.html* con URI <https://cercador.dlsi.ua.es/tt/X/ttpro.html> (donde *X* es el NIF o el NIE sin la letra final), accesible a través de un enlace desde el documento maestro.
- El documento *ttpro.html* es (X)HTML válido.
- El documento *ttpro.html* enlaza:
 - los ficheros originales.
 - los ficheros traducidos antes de hacer el control de calidad.
 - la memoria de traducción generada por OmegaT como resultado del proceso de traducción antes de hacer el control de calidad.
 - cualquier recurso que hayáis usado para traducir: traducciones automáticas en bruto, memorias de traducción, glosarios que hayáis creado, etc.
 - los ficheros traducidos después de hacer el control de calidad.
 - la memoria de traducción generada por OmegaT con las correcciones que hayáis hecho durante el control de calidad.
 - tres ficheros en formato de texto plano (*.txt*): uno con el texto original usado para la evaluación de la traducción automática, otro con la traducción en bruto y otro con la traducción posteditada.
- Los documentos traducidos están en el formato especificado en este enunciado y en cada proyecto de traducción.

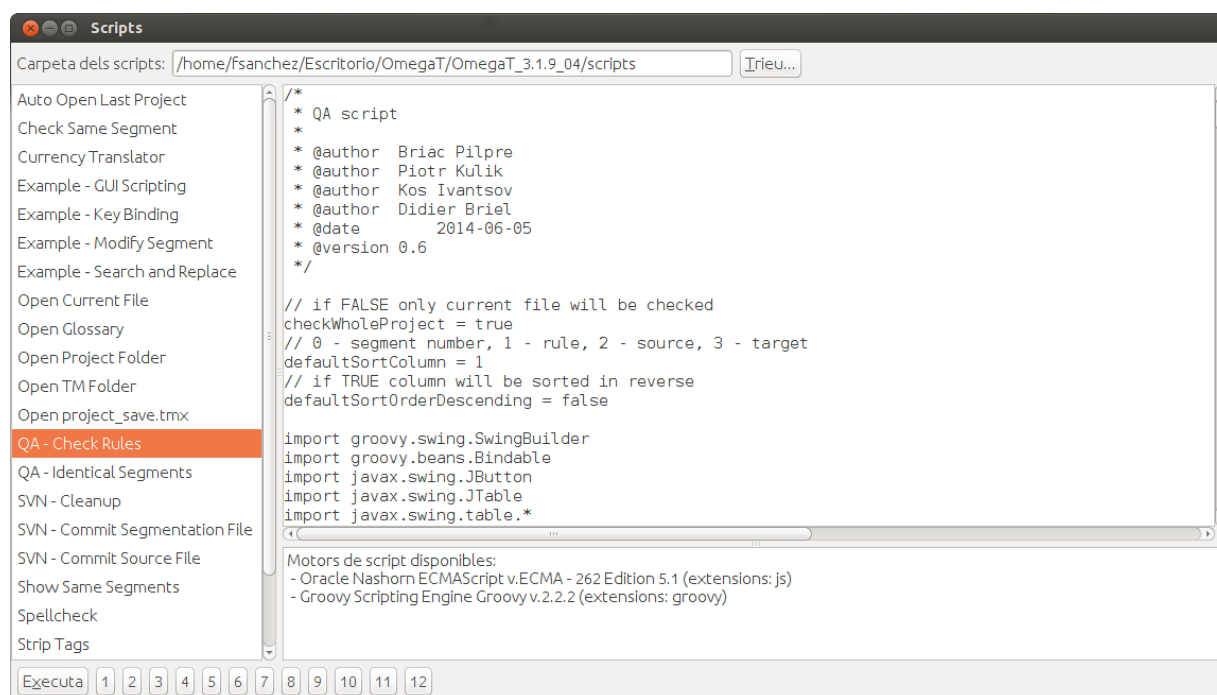


Figura 1: Pantalla con los programas de OmegaT donde se muestra seleccionado el programa QA – Check Rules.

Aspectos del proyecto que se valorarán:

- Los documentos del proyecto están bien traducidos y en el mismo formato que los originales.
- La redacción del informe `ttpro.html` es adecuada y recoge toda la información mencionada en la plantilla de informe.
- El informe describe de forma adecuada el proyecto de traducción que se ha llevado a cabo.
- En el informe se explican claramente las decisiones de adopción de las tecnologías de la traducción que se han adoptado y las razones. Estas decisiones deben estar basadas en criterios *objetivos*.
- Cuando en el informe se describe la metodología para hacer alguna de las tareas del proyecto, la descripción permite que una persona ajena a la asignatura y sin acceso a este enunciado pueda reproducir vuestro trabajo.
- Cuando se facilitan datos, como por ejemplo el coste de traducción por palabra, se da una indicación del margen de variación de los valores referidos.

A. Recursos

Para llevar a cabo vuestro proyecto de traducción podéis usar cualquier recurso que os sea de ayuda además de los proporcionados por el profesorado para cada proyecto concreto. He aquí una lista no exhaustiva de recursos que pueden ser de vuestro interés:

- Base de datos terminológica de la Unión Europea: <http://iate.europa.eu>
- Base de datos terminológica del Termcat: <http://www.termcat.cat>
- El servicio de búsqueda de memorias de traducción *My Memory Translated* (<http://mymemory.translated.net>) permite obtener memorias de traducción a partir de los ficheros a traducir. Este servicio requiere que os registréis para poder descargar memorias de traducción en formato TMX.
- Los buscadores de traducciones <http://www.linguee.es> y <http://context.reverso.net/translation/> permiten buscar la traducción de una palabra o expresión en

otro idioma. Estos buscadores hacen la búsqueda en textos paralelos bilingües.

- Los servicios de alineamiento de traducciones de YouAlign (<http://www.youalign.com>) o de Prompsit (<http://aplica.prompsit.com/es/tmx/>) que podéis usar para alinear un texto con su traducción automática y obtener un fichero TMX para usar con OmegaT (véase el apéndice B).
- Las memorias de traducción del proyecto OPUS (Open Parallel Corpus): <http://opus.lingfil.uu.se/>

Además de los recursos mencionados más arriba podéis usar cualquiera otro recurso o traductor automático de los que hay disponibles en Internet.

B. Uso de traducción automática desde OmegaT

Para usar uno de los traductores automáticos soportados por OmegaT tenéis que seleccionar el traductor automático a utilizar desde el menú Opciones→Traducción automática. Mientras traducís tenéis que usar la combinación de teclas Ctrl-M para reemplazar el segmento que estáis traduciendo por su traducción automática; traducción que a continuación tenéis que posteditar.

B.1. Google Translate

Para usar el traductor de Google, se tiene que pagar (20 dólares por cada millón de caracteres) y obtener una clave que se tiene que instalar en OmegaT. Cuando tengáis la clave tenéis que incluirla en el fichero OmegaT.bat (Windows) u OmegaT (GNU/Linux y MacOS) que estará en la carpeta donde descomprimisteis OmegaT. Para esto abrid con la *Libreta* o *Bloc de notas* el fichero y añadid el texto `-Dgoogle.api.key="valor-de-la-clave"` después de la palabra `java`. La línea completa tiene que quedar como se indica a continuación: `java -Dgoogle.api.key="valor-de-la-clave" -jar -Xmx1024M OmegaT.jar %*`.

Para que los cambios tengan efecto tenéis que iniciar OmegaT haciendo siempre doble clic sobre el fichero OmegaT.bat (en lugar de OmegaT.jar). Si vuestro ordenador tiene un sistema operativo GNU/Linux o MacOS, tenéis que hacer lo mismo pero con el fichero OmegaT.

Para más información consultad la URI http://www.omegat.org/en/howtos/google_translate.html.

B.2. Bing Translator

En el caso de Bing Translator, es necesario darse de alta en <https://datamarket.azure.com/dataset/bing/microsofttranslator> y obtener dos claves que se tienen que instalar en OmegaT. Cuando os hayáis dado de alta, para obtener las claves tenéis que iniciar sesión e ir a la opción Mi cuenta. Los datos que necesitáis son la clave de cuenta principal y el identificador del cliente. Estos datos los tenéis que incluir en el fichero OmegaT.bat (Windows) u OmegaT (GNU/Linux, MacOS), que estará en la carpeta donde descomprimisteis OmegaT. Para añadir las claves abrid con *Notepad++* el fichero y añadid el texto `-Dmicrosoft.api.client_secret="valor-de-la-clave-principal"` `-Dmicrosoft.api.client_id="identificador-del-cliente"` después de la palabra `java`. La línea completa debe quedar como se indica a continuación: `java -Dmicrosoft.api.client_secret="valor-de-la-clave-principal" -Dmicrosoft.api.client_id="identificador-del-cliente" -jar -Xmx1024M OmegaT.jar %*`.

Como antes, para que los cambios tengan efecto tenéis que iniciar OmegaT haciendo doble clic sobre el fichero OmegaT.bat (Windows) u OmegaT (GNU/Linux y MacOS).

Para más información consultad la URI http://www.omegat.org/en/howtos/ms_translator.html.

B.3. Apertium

En el caso de Apertium, os recomendamos instalar el plugin para OmegaT que podéis descargar desde <https://apertium.svn.sf.net/svnroot/apertium/builds/apertium-omegat/apertium-omegat.jar>. Para instalarlo sólo tenéis que copiar el archivo `apertium-omegat.jar` en la carpeta `plugin` de vuestra instalación de OmegaT. Podéis encontrar más información en <http://wiki.apertium.org/wiki/apertium-omegat>.

B.4. Otros traductores

Para usar un traductor automático que no esté soportado directamente por OmegaT hay que obtener un fichero TMX que en lugar de contener una memoria de traducción contendrá los segmentos en lengua origen a traducir y su traducción automática. Este fichero se puede obtener fácilmente siguiendo los pasos siguientes:

1. Abrid el documento a traducir con OmegaT, seleccionad todo el texto (desde dentro de OmegaT) y copiadlo.
2. Cread un documento de texto plano (`origen.txt`) nuevo desde el Bloc de Notas (o la Libreta) y pegar el texto que habéis copiado en el paso anterior.
3. Abrid el navegador y pegar el texto en el traductor automático que deseéis utilizar y traducirlo.
4. Seleccionad todo el texto traducido automáticamente y copiadlo.
5. Cread un documento de texto plano (`traduccion.txt`) nuevo desde el Bloc de Notas (o la Libreta) y pegar el texto traducido que habéis copiado en el paso anterior.
6. Id a uno de los alineadores de documentos en línea, <http://www.youalign.com> o <http://aplica.prompsit.com/es/tmx/> y alinead los textos creados en los pasos 2 y 5. Como resultado obtendréis una memoria de traducción en formato TMX. Dad a este fichero un nombre que os permita identificar que se trata de traducciones automáticas (por ejemplo `trad-auto-nombre-traductor.tmx`).

El fichero TMX resultante lo tenéis que copiar a la carpeta `tm` de vuestro proyecto de OmegaT. Cuando estéis

traduciendo OmegaT os ofrecerá siempre una propuesta de traducción con porcentaje de coincidencia del 100 %; tened en cuenta que se trata de una traducción automática (lo sabréis por el nombre del fichero del que viene la unidad de traducción) que tenéis que posteditar.

Advertencia: Si vais a usar las estadísticas de coincidencias para calcular los costes del proyecto de traducción (apartado 4.4), aseguraos que en la carpeta `tm` de vuestro proyecto de traducción no hay ninguna memoria de traducción que contenga traducciones automáticas cuando calculéis estas estadísticas para no falsear los datos.

C. Solución de problemas con el proyecto en grupo de OmegaT

Si tenéis **problemas de sincronización** del proyecto de traducción tenéis que hacer lo siguiente para solucionarlos:

1. Cerrad OmegaT.
2. Haced un copia de seguridad de la carpeta del proyecto por si acaso.
3. Instalad TortoiseSVN (<https://tortoisesvn.net>), si no lo tenéis ya instalado.
4. Id a la carpeta del proyecto de traducción.
5. Seleccionad la carpeta y haced clic con el botón derecho del ratón.
6. En el menú desplegable que aparece, seleccionad la opción TortoiseSVN→Clean Up.
7. Marcad la opción “remove locks” (o “break locks”) y aceptad.
8. Haced clic con el botón derecho del ratón nuevamente sobre la carpeta del proyecto y seleccionad la opción SVN Update.
9. Haced clic con el botón derecho del ratón un vez más sobre la carpeta del proyecto y seleccionad la opción SVN Commit.

Si todo ha ido bien el proyecto tendría que volver a estar sincronizado. Podéis volver a abrirlo desde OmegaT para comprobarlo.